

Naval Research Laboratory

Washington, DC 20375-5320



NRL/FR/5740--99-9918

A Fuzzy Logic Multisensor Association Algorithm: Multiple Emitters, Computational Complexity, and Noisy Data

JAMES F. SMITH III

*Surface Electronic Warfare Systems Branch
Tactical Electronic Warfare Division*

July 15, 1999

19990831 162

Approved for public release; distribution is unlimited.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188
<p>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.</p>			
1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED	
	July 15, 1999	Final Report 10/1/97-9/30/98	
4. TITLE AND SUBTITLE A Fuzzy Logic Multisensor Association Algorithm: Multiple Emitters, Computational Complexity, and Noisy Data			5. FUNDING NUMBERS
6. AUTHOR(S) James F. Smith III			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Research Laboratory Washington, DC 20375-5320			8. PERFORMING ORGANIZATION REPORT NUMBER NRL/FR/5740--99-9918
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Office of Naval Research 800 North Quincy Street Arlington, VA 22217-5660			10. SPONSORING/MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES			
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited.			12b. DISTRIBUTION CODE
13. ABSTRACT (Maximum 200 words) A recursive multisensor association algorithm based on fuzzy logic has been developed. It simultaneously determines fuzzy grades of membership and fuzzy cluster centers. It is capable of associating data from various sensor types. In its simplest form, it makes no assumption about noise statistics as many association algorithms do. The algorithm is capable of performing without operator intervention. It associates data from the same target for multiple sensor types. The algorithm also provides an estimate of the number of targets present, reduced noise estimates of the quantities being measured, and a measure of confidence to assign to the data association. A comparison of the algorithm to a more conventional Bayesian association algorithm is provided. Data from both the electronic support measures (ESM) and radar systems are noisy, and ESM data are also intermittent. The data has probability of detection less than unity. The effect of a large number of targets being present in the data on parameter estimation, determination of the number of targets, and multisensor data association is examined. Finally, issues related to computational complexity are discussed.			
14. SUBJECT TERMS Fuzzy clustering Fuzzy sets Fuzzy logic			15. NUMBER OF PAGES 21
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL

CONTENTS

1. INTRODUCTION.....	1
2. FUZZY SETS, CLUSTERING, DEFUZZIFICATION, AND SUPERCLUSTERING	1
2.1 Fuzzy Set Theory	1
2.2 Fuzzy Clustering	2
2.3 Superclustering	3
3. THE TRUNK-WILSON ALGORITHM	4
4. PROBABILITY AUGMENTED SUPERCLUSTERING	6
5. APPLICATION OF THE FUZZY ASSOCIATION ALGORITHM TO SIMULATED DATA AND COMPARISON TO THE TW ALGORITHM	6
5.1 Association of ESM and Noiseless Radar	6
5.2 Association of Noisy ESM and Noisy Radar Measurements	8
5.3 Association of Data with Random Dropout.....	9
5.4 ESM and Radar Association for Multiple Targets	10
5.5 Multiple Closely Spaced Emitters	12
6. SLIDING WINDOWS, PERFORMANCE, AND COMPUTATIONAL COMPLEXITY	14
7. FUTURE APPLICATIONS	16
8. CONCLUSIONS	17
9. ACKNOWLEDGMENTS	17
REFERENCES	17

A FUZZY LOGIC MULTISENSOR ASSOCIATION ALGORITHM: MULTIPLE EMITTERS, COMPUTATIONAL COMPLEXITY, AND NOISY DATA

1. INTRODUCTION

The problem considered in this report is how to associate electronic support measures (ESM) signals with one or more of m possible radar tracks. An algorithm based on fuzzy set theory has been developed to solve this problem. It considers the complexities offered by having multiple radar tracks and unequal numbers of measurements. It is capable of making its own estimate from ESM data of bearing. As such, it provides additional measures of association, unlike the Trunk-Wilson (TW) Bayesian theory [1] with which it is compared. This algorithm can estimate the number of targets present in the data and use fuzzy set theoretic techniques to suppress outliers. The fuzzy grades of membership provide opportunities for incorporation of heuristic rule sets and extension to probability theory. The fuzzy cluster centers represent reduced noise estimates of the measured quantities. Finally, in comparison to an existing Bayesian algorithm the fuzzy association algorithm exhibits superior performance.

Section 2 introduces the concepts of fuzzy set theory, hard and fuzzy clustering, defuzzification, and superclustering. Section 3 introduces the TW algorithm, an established algorithm with which the fuzzy algorithm will be compared. Section 4 discusses how, in a recursive algorithm, the process of estimating the number of targets present in the data through superclustering can be improved by introducing a priori information. Section 5 discusses application of the fuzzy association algorithms to simulated ESM and radar data. In section 5, both noiseless and noisy radar measurements are considered in different examples. The effects of random data loss and also multiple targets present in the data are examined. Section 6 discusses the algorithm's computational complexity. In section 7, the fuzzy association algorithm's potential for incorporation into a suite of fuzzy algorithms for resource allocation management of a multifunction antenna is discussed. Finally, section 8 provides conclusions.

2. FUZZY SETS, CLUSTERING, DEFUZZIFICATION, AND SUPERCLUSTERING

The development of the fuzzy association algorithm requires the concepts of the fuzzy set, clustering, fuzzy clustering, defuzzification, and superclustering. These concepts are developed in the following subsections.

2.1 Fuzzy Set Theory

This section provides a basic introduction to the ideas of fuzzy set theory. Fuzzy set theory allows an object to have partial membership in more than one set. It does this through the introduction of a function known as the membership function, which maps from the complete set of objects X into a set known as membership space. More formally the definition of a fuzzy set [2] is

If \mathbf{X} is a collection of objects denoted generically by x , then a fuzzy set A in \mathbf{X} is a set of ordered pairs:

$$A = \{(x, \mu_A(x)) \mid x \in X\}.$$

$\mu_A(x)$ is called the membership function or grade of membership (also degree of compatibility or degree of truth) of x in A , which maps X to the membership space M . (When M contains only the two points 0 and 1, A is nonfuzzy and $\mu_A(x)$ is identical to the characteristic function of a nonfuzzy set.) The range of the membership function is a subset of the non-negative real numbers whose supremum is finite. Elements with a zero degree of membership are normally not listed.

2.2 Fuzzy Clustering

Fundamental to the development of the fuzzy association algorithm is the concept of clustering. Clustering is an operation that allows data to be grouped into classes defined by a similarity measure. By definition [3], given K objects the algorithm forms N clusters such that, with respect to the similarity measure, the members of each cluster have a greater similarity to each other than to the members of any other cluster.

The kind of clustering used for association is known as fuzzy clustering. Reference 4 describes a batch version of the algorithm. There are many kinds of clustering. If each point is assigned 100% to a particular cluster, the algorithm is referred to as a hard clustering algorithm. Fuzzy clustering differs from hard clustering in that each data point can have partial assignment in each cluster. The grade of membership previously defined under the concept of a fuzzy set gives the degree of membership of each point in each cluster. The particular type of fuzzy clustering used here is referred to as a c-means algorithm [5,6].

Clustering plays a significant role in the development of the fuzzy association algorithm. The grades of membership are established by minimizing a functional. This functional is found in many places in the literature of fuzzy sets and fuzzy clustering [5,6]. It is defined below after some preliminary notation is established.

Let X be any finite set; V_{cn} is the set of real $c \times n$ matrices; c is an integer with $2 \leq c \leq n$, and n is the number of data points. The fuzzy c -partition space for X is the set

$$M_{fc} = \{U \in V_{cn} \mid u_{ik} \in [0,1] \forall i, k; \sum_{i=1}^c u_{ik} = 1 \forall k; 0 < \sum_{k=1}^n u_{ik} < n \forall i\}.$$

Row i of a matrix $U \in M_{fc}$ exhibits (values of) the i th membership function (or i th fuzzy subset) U_i in the fuzzy c -partition U of X . Stated less formally, u_{ij} is the grade of membership of data point j in fuzzy cluster i .

Definition: Let $J_m: M_{fc} \times R^{cp} \rightarrow R^+$,

$$J_m(U, v) = \sum_{k=1}^n \sum_{i=1}^c (u_{ik})^m (d_{ik})^2,$$

where R^{cp} is the collection of possible p -dimensional vectors with real elements taken c at a time, and R^+ is the real interval $[0, \infty)$;

$$U \in M_{fc}$$

is a fuzzy c -partition of X ;

$$v = (v_1, v_2, \dots, v_c) \in R^{cp} \quad \text{with } v_i \in R^p$$

is the cluster center or prototype of u_i , $1 \leq i \leq c$;

$$(d_{ik})^2 = \|x_k - v_i\|^2 \text{ and } \|\cdot\|$$

is any inner product induced norm on R^p ; and weighting exponent $m \in [1, \infty)$.

Since each term of J_m is proportional to $(d_{ik})^2$, J_m is a square-error clustering criterion. The solution of the fuzzy clustering problem consists of minimizing J_m as a function of U and v subject to the constraints imposed in the definition of M_{fc} .

The value of m used for the simulations in section 5 is $m = 2$. The significance of the m -parameter is more fully discussed elsewhere [5,6].

The goal of the fuzzy clustering algorithm is to determine fuzzy cluster centers v_i that represent the average value of the quantities in the fuzzy clusters, and the grade of membership of the k th data point in the i th fuzzy cluster for all data points- k and clusters- i . The algorithm determines these quantities by minimizing a least-square cost function in which each term is weighted by a power of the grade of membership. Each term of the cost function simultaneously measures the distance of the data point from a cluster center and is weighted by the point's membership in that cluster. The minimization is conducted subject to the constraints that the sum of the grades of membership over clusters for a particular data point must equal unity, and for each cluster the sum of grades of membership over data points must be bound between 1 and the maximum number of data points.

The system of equations resulting from the minimization represents a coupling between the fuzzy cluster centers and the fuzzy partition matrix. An initial estimate of either quantity is all that is required to initialize the iteration process. Thus, an initial estimate of the fuzzy cluster center by one class of sensors can be used to cluster data measured by another sensor system. Thus the algorithm can be made recursive by using previous estimates of cluster centers and/or cluster center estimates derived from other sensors.

The output consists of:

- high-quality estimates of the grades of membership; these quantities provide a measure of confidence of how well the data are clustered and a means of making an optimal data-point cluster assignment. This is especially useful if a data point falls on the boundary between clusters; and
- fuzzy cluster centers, which represent reduced noise values of the measured quantities. Cluster centers are also useful for conveying the position of the cluster.

For many applications, it is necessary to extract from the clustering algorithm a non-fuzzy, i.e., crisp, statement of the assignment of each point. The process of taking fuzzy results and extracting definite, i.e., crisp, data-point cluster assignment is known as *defuzzification*.

The current approach to defuzzification consists of making a definite data point assignment to that cluster for which the data point has the largest grade of membership. If a data point has equal grades of membership for more than one cluster, the point, in the simplest form of defuzzification, is assigned to the first cluster that is encountered.

2.3 Superclustering

Clustering algorithms, including the fuzzy clustering algorithm, generally require a specification of the final number of clusters. If the data being clustered represent ships, aircraft, missiles, etc., this implies a priori knowledge of the number of targets. Obviously, the number of targets will not be known in general before processing. So it is desirable to develop a technique for determining from the data the appropriate number of clusters, i.e., the number of targets. Such a technique, known as *superclustering*, has been developed which provides a solution to this problem. The superclustering techniques developed here are related to and represent an extension of techniques in fuzzy cluster validity theory [6].

The method of superclustering is described in greater detail in the literature [4,7] and summarized here. Given an upper bound on the number of clusters, this bound is supplied to the fuzzy clustering algorithm. The fuzzy clustering algorithm produces this number of clusters for the data, with associated grades of membership for each data point in each cluster. The fuzzy clustering algorithm also provides the coordinates of the fuzzy cluster centers. Intuitively, clusters should be separated, nonoverlapping, and not extremely close to each other with respect to some measure. It then becomes essential to define a measure of “closeness” and provide a criterion for what “too close” means.

An obvious candidate for a measure of closeness of two clusters is the separation of the cluster centers. The cluster centers do not tell the whole story. The data points may be distributed close to the cluster center or they may be a significant absolute distance from it. Also, when dealing with fuzzy clustering, before defuzzification, the points generally do not belong 100% to any cluster. To provide a unitless measure of closeness and incorporate the concept of vagueness inherent in fuzzy algorithms, the distance between fuzzy cluster centers should be normalized by some function of the grades of membership. Incorporation of the grades of membership, i.e., superclustering, before defuzzification has the advantage of potentially better cluster assignments for points that fall on the boundary between clusters.

One such normalized measure of cluster center separation is the c-matrix defined below. Let $v(i)$ and $v(j)$ be the position vectors for the fuzzy cluster centers for cluster i and cluster j , respectively, and N the number of data points. Then the i th – j th element of the c-matrix is

$$c(i, j) = \|v(i) - v(j)\| / \max(\text{std}(i), \text{std}(j)), \quad (1)$$

where

$$\text{std}(k) = \sqrt{\sum_{i=1}^N u(i, k)^m * (x(i) - \text{mean}(k))^2 / \sum_{i=1}^N u(i, k)^m} \quad (2)$$

and

$$\text{mean}(k) = \left(\sum_{i=1}^N u(i, k)^m * x(i) / \sum_{i=1}^N u(i, k)^m \right). \quad (3)$$

Equations (2) and (3) define the fuzzy standard deviation and the fuzzy mean, respectively.

The c-matrix capitalizes on the intuitive idea that cluster centers should be separated by a certain number of fuzzy standard deviations. If cluster centers are closer than this, they probably correspond to the same cluster. If it is determined that two or more clusters should be merged into a single cluster, the resulting grouping will be referred to as a *supercluster*. A criterion must be established to determine when supercluster formation is warranted. A simple criterion consists of defining a threshold τ , such that if $c(i,j) < \tau$, then clusters i and j are merged into a supercluster. References 4 and 7 discuss methods for selecting the value of τ .

The next step involved in superclustering after c-matrix formation and establishment of the threshold is determining exactly how to form superclusters, i.e., when there is more than one choice based on what has been developed up to now, what is the best supercluster formation scheme. The formation schemes are described in greater detail in Refs. 4 and 7.

3. THE TRUNK-WILSON ALGORITHM

The TW algorithm assumes there are K ESM tracks, each specified by a different number of ESM measurements. These ESM measurements are associated with either no radar track or one of m radar tracks,

each radar track having a different number of measurements. The association of ESM tracks with radar tracks using the multiple hypothesis testing technique is as follows:

H_0 : ESM measurements are associated with no radar track

H_1 : ESM measurements are associated with the first radar track

H_j : ESM measurements are associated with the j th radar track

H_m : ESM measurements are associated with the m th radar track.

The Bayesian procedure, which minimizes the probability of error, is to select the hypothesis having the largest a posteriori probability. It is assumed that the ESM measurement errors are independent and Gaussian distributed, with zero mean and constant variance σ^2 . If the a priori probabilities are equal, the minimum error decision rule selects the target j based on the statistical distance d_j for which d_j is minimized and given by

$$d_j = \sum_{i=1}^{K_j} [\theta_e(t_i) - \theta_j(t_i)]^2 / \sigma^2 \quad j = 1, \dots, m, \quad (4)$$

where $\{\theta_e(t_i), 1 \leq i \leq K_j\}$ is a set of K_j ESM bearing measurements, and $\theta_j(t_i)$ is the true bearing to target j at the time of the i th ESM measurement. The minimized statistical distance $d_{j\min}$ has a chi-square density with K_j degrees of freedom. Consequently, the desired a posteriori probability P_j is given by

$$P_j = \int_{d_j}^{\infty} \chi^2(t) dt, \quad (5)$$

where $\chi^2(t)$ is the chi-square density function. The TW algorithm selects the association that has the largest probability, P_{\max} where

$$P_{\max} = \max \{P_j \mid j = 1, \dots, m\}.$$

The TW algorithm also uses the next largest probability P_{next} and four decision theoretic quantities obtained by Trunk and Wilson. Three of the quantities are the high (T_H), middle (T_M), and low (T_L) probability thresholds. The fourth quantity is the probability margin (R). The corresponding decision rules are:

- (1) firm correlation, $P_{\max} \geq T_H$ and $P_{\max} \geq P_{\text{next}} + R$ — the ESM signal goes with the radar track having largest P_j (i.e., P_{\max}),
- (2) tentative correlation, $T_H > P_{\max} \geq T_M$ and $P_{\max} \geq P_{\text{next}} + R$ — ESM signal probably goes with radar track having largest P_j (i.e., P_{\max}),
- (3) tentative correlation with some track, $P_{\max} \geq T_M$ but $P_{\max} < P_{\text{next}} + R$ — ESM signal probably goes with some radar tracks (but the algorithm cannot determine which),
- (4) tentative uncorrelated, $T_M > P_{\max} \geq T_L$ — ESM signal probably does not go with any radar track, and
- (5) firmly uncorrelated, $T_L \geq P_{\max}$ — ESM signal does not go with any radar track.

The threshold T_H is set equal to P_{FA} , defined as the probability of falsely associating a radar track with an ESM signal when the ESM signal does not belong with the radar track. The threshold T_H is a function of the

azimuthal difference, denoted by μ , between the true (ESM) position and the radar track under consideration. The threshold T_H was determined for two values $\mu = 1.0 \sigma$ and 1.5σ by Trunk and Wilson using simulation techniques. Also, the threshold T_M was determined using the simulation previously used to determine T_H , for the same two values of μ , 1.0σ and 1.5σ . The threshold T_L is defined as a rejection rate of P_R . In this report, $\mu = 1.5 \sigma$ is used and the threshold T_L is set equal to 0.001.

The probability margin R is determined by specifying a probability of an association error P_e according to

$$P_e = P_R \{P_{\max} \geq P_{\text{next}} + R\},$$

where P_{\max} corresponds to an incorrect association and P_{next} corresponds to the correct association. The probability margin R is a function of P_e and the separation μ of the radar tracks. The values of R were determined by Trunk and Wilson for $P_e = 0.01$ by simulation techniques.

4. PROBABILITY AUGMENTED SUPERCLUSTERING

In many cases, superclustering alone can determine the number of targets exactly or within one target if there are enough data [4]. In a recursive algorithm, data are arriving from one moment to the next. Adding in a priori information, e.g., noise statistics, can reduce the amount of data required to make a correct decision as to the number of targets. By calculating the probability that the data are associated with each cluster center, cluster centers with a probability of association much less than the maximum can be neglected. The rejection threshold that has been found to be useful is 20% of the maximum probability. This procedure is referred to as *probability augmented superclustering* and has been found to be quite effective [4].

5. APPLICATION OF THE FUZZY ASSOCIATION ALGORITHM TO SIMULATED DATA AND COMPARISON TO THE TW ALGORITHM

In this section, the fuzzy association algorithm is examined for five different kinds of test cases and compared to the TW algorithm. In subsection 5.1 there is one target with constant bearing of 0° . The ESM measurements are simulated by adding 1° standard deviation Gaussian noise to the true bearing 0° . There are two radar cases, each with three noiseless radar measurements. All ESM and radar points are detected. Subsection 5.2 differs from 5.1 only in that 0.1° standard deviation Gaussian noise is added to the radar measurements. Subsection 5.3 adds the additional feature that not all radar points are detected. In subsection 5.4, the radar is noisy and all points are detected, but there are four to 10 emitters. Finally in subsection 5.5, both the fuzzy association algorithm and TW algorithm are applied to the extremely difficult example of 10 emitters very closely spaced in the RF-PRI plane.

5.1 Association of ESM and Noiseless Radar

Two simulation classes are considered here. In each case, the target has a constant bearing of 0° , and there are three constant bearing radar estimates. The radar estimates are what provide the distinction between the four examples:

Example 1: $\mu_1 = 0^\circ$, $\mu_2 = 1^\circ$, $\mu_3 = -1^\circ$;

Example 2: $\mu_1 = 2^\circ$, $\mu_2 = 1^\circ$, $\mu_3 = -1^\circ$.

In each simulation, two data points are added each time until total of 50 data points are accumulated. In all four simulations, ensemble averages have been conducted.

In section 3, five hypothesis classes are defined, but results for only two are given in this paper. The remaining hypothesis classes are generally of low probability, carry little information, and tend to obscure the most important results. Firm correlation between the ESM data and the 0° radar track is denoted by **FCT** throughout the text and figures. Likewise, when the association algorithms determine that the ESM data are firmly not associated with any of the radar tracks, this is denoted by **FNT**.

The fuzzy association algorithm compares its estimate of bearing to the radar measurements and also uses the fuzzy standard deviation for σ . So when Eq. (4) is used in the fuzzy association algorithm,

$\theta_e(t_i)$ is the fuzzy bearing estimate at the time t_i

and

σ is the fuzzy standard deviation.

The substitution of the fuzzy bearing estimate and the fuzzy standard deviation into Eq. (4) is what distinguishes the fuzzy association algorithm from the TW association algorithm. All definitions of correlation remain the same.

Figure 1 presents results for the radar example $\mu = 0^\circ, 1^\circ, -1^\circ$. Since the radar results contain truth, i.e., a target moving with constant bearing of 0° , a good association algorithm will establish that there is a firm correlation between the ESM data and the 0° bearing track. Figure 1 plots the probability that the association algorithms establish a firm association between ESM data and the radar measurements. The fuzzy association algorithm results are given by the curve marked with o's and the TW results are indicated by the curve marked with +'s. The vertical axis indicates probability of firm correlation with the 0° radar track. The horizontal axis indicates the number of data points necessary to establish that level of probability.

The fuzzy association algorithm results are always superior to the TW algorithm. At 10 data points the fuzzy algorithm has established a 70% probability of firm correlation, whereas the TW algorithm requires about 26 points to establish the same level of probability of FCT. The fuzzy algorithm establishes an 80% probability of FCT by the 13th data point, whereas the TW algorithm requires about 32 points to reach the same level of success. The fuzzy algorithm reaches 90% probability of FCT at 18 data points and the TW algorithm at about the 39th point. So the fuzzy algorithm establishes high probabilities of firm correlation with 1/3 to 1/2 the data required by the TW algorithm. In this sense, the fuzzy algorithm is 2 to 3 times faster

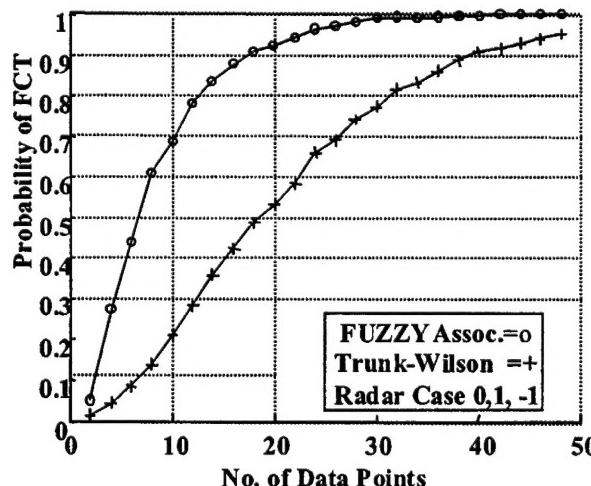


Fig. 1 — Perfect radar for probability of FCT

than the TW algorithm. Also, this is a difficult example for any association algorithm since there are two additional radar measurements within one noise standard deviation. The data window accommodated up to 50 points in this case, although a significantly smaller sliding window could be used. The ability of the fuzzy algorithm to make high quality decisions with much less data than the TW algorithm is significant since real data are frequently sparse and intermittent.

Figure 2 presents results for the radar example $\mu = 2^\circ, 1^\circ, -1^\circ$. Since the radar results do not contain truth, i.e., a target moving with constant bearing of 0° , a good association algorithm will establish that the ESM data are firmly not correlated with the radar tracks. Figure 2 plots the probability the association algorithms establish that the ESM and radar data are firmly not correlated. By the eighth data point, the fuzzy algorithm has reached a 55% probability of FNT, whereas the TW algorithm never exceeds that probability. For this example, the fuzzy algorithm is 6 times faster than the TW algorithm, i.e., it reaches the TW algorithm's maximum probability of FNT with $1/6$ the data. The fuzzy algorithm has 80 and 90% probability of FNT by the 13th and 20th data points, respectively. Thus once again, the fuzzy algorithm makes a high-quality decision long before the TW algorithm.

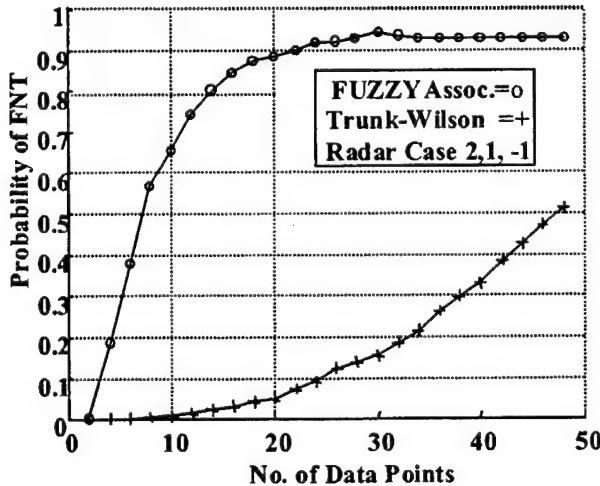


Fig. 2 — Perfect radar for probability of FNT

5.2 Association of Noisy ESM and Noisy Radar Measurements

In the previous simulations it was assumed that the radar measurements were noiseless. The TW algorithm can be used to associate noisy ESM and noisy radar measurements [8] as follows. The radar measurements for radar track j at time t_i will have zero mean Gaussian noise added to them. The variance of the noise will be denoted as σ_{ij}^2 for the j th radar track and the i th time. If the variance in Eq. (4) is replaced by

$$\sigma^2 = \sigma_E^2 + \sigma_{ij}^2,$$

where σ_E^2 = the variance of the ESM noise, then the statistic defined in Eq. (4) still has a chi-square density. Thus, it follows that all the thresholds should have the same value, whether or not the radar measurements are noisy. There will be some difference in multitarget performance because of the different dependence between the squared errors $\{d_j, j = 1, 2, \dots, m\}$ due to the radar variances.

Figure 3 is for the radar example $\mu = 0^\circ, 1^\circ, -1^\circ$ with $\sigma_{ij} = 0.1^\circ$ for all times t_i and radar tracks j . This figure corresponds to the perfect radar case given in Fig. 1. The radar noise standard deviation is consistent with levels found in modern radar systems. Figures 1 and 3 are practically indistinguishable. This implies that both algorithms are insensitive to the small amounts of noise found in modern radar systems.

Figure 4 corresponds to the radar example $\mu = 2^\circ, 1^\circ, -1^\circ$. Once again, the radar noise has a standard deviation of $\sigma_{ij} = 0.1^\circ$ for all times t_i and radar tracks j . Figure 4 corresponds to the noiseless cases provided in Fig. 2. Once again, both the fuzzy association and TW association algorithms show little sensitivity to the radar noise.

5.3 Association of Data with Random Dropout

In this subsection, ESM and radar data are associated for a single emitter and three noisy radar measurements with the same noise standard deviation as in subsection 5.2. Unlike the results in 5.2, there have been random dropouts of both ESM and radar data points. In Figs. 5 and 6, radar results for example 1 have been used, i.e., $0, -1, 1$, with 0.1° standard deviation Gaussian noise added. The vertical axis always measures the probability of FCT and the horizontal axis time in microseconds (actually, time is in arbitrary units). The horizontal axes in Figs. 1-4 are labeled "No. of Data Points," but the time allowed for data point accumulation is the same as in Figs. 5 and 6. If the horizontal axes are thought of as measured in units of time, then direct comparison between Figs. 1 to 4 and Figs. 5 and 6 is possible. The results from the fuzzy association algorithm are marked with o's and the Trunk-Wilson algorithm by +'s.

In Fig. 5, the probability of detection of ESM points is 90%. In this case as observed before, the fuzzy algorithm reaches a 70% probability of FCT 2.6 times faster than the TW algorithm and 90% probability of FCT a little more than 2 times faster than the TW algorithm. The reduction in probability of detection of the ESM data by 10% has almost no effect on the fuzzy association algorithm, but the TW algorithm is showing deterioration.

In Fig. 6, the probability of detection of ESM points is 70%. In this case, as observed before, the fuzzy algorithm reaches 70 and 80% probability of FCT two times faster than the TW algorithm. The fuzzy association algorithm reaches 90% probability of FCT shortly after reaching 80% probability of FCT, but the TW algorithm never exceeds 86% probability of FCT in this time window.

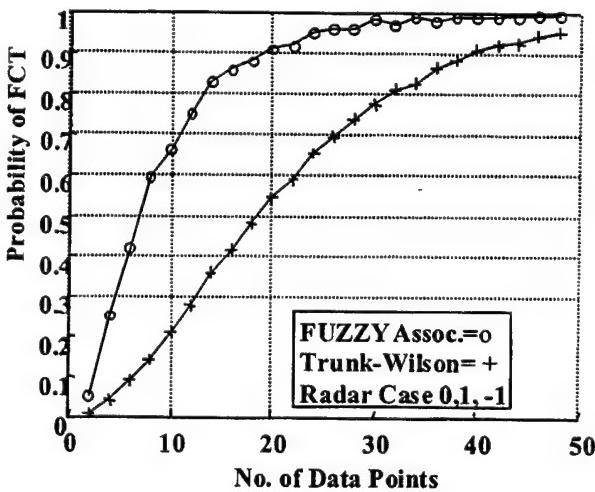


Fig. 3 — Noisy radar for probability of FCT

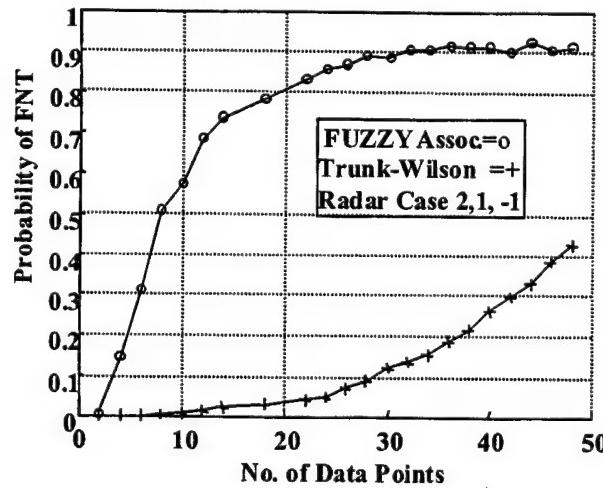


Fig. 4 — Noisy radar for probability of FNT

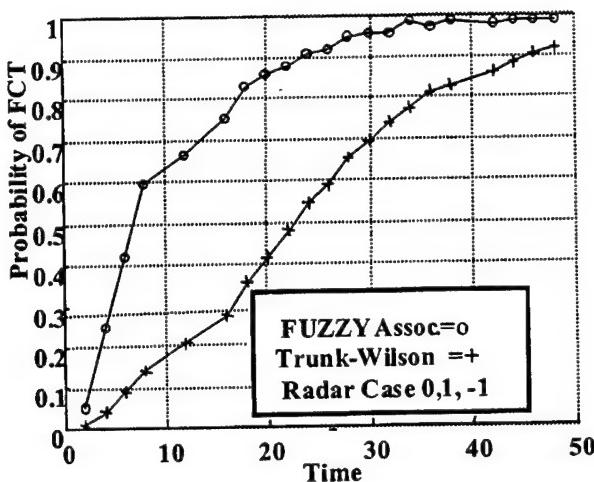


Fig. 5 — 10% random data loss and noisy radar for probability of FCT

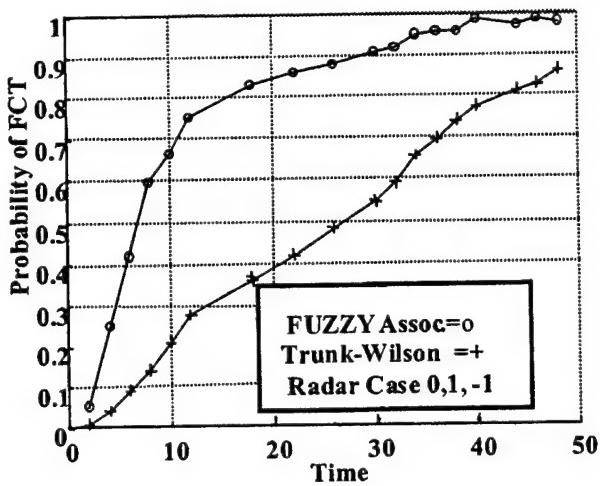


Fig. 6 — 30% random data loss and noisy radar for probability of FCT

5.4 ESM and Radar Association for Multiple Targets

In this section, the data contain measurements for multiple emitters. The 4-emitter (Figs. 7, 9, and 10) and 10-emitter cases (Figs. 8, 11, and 12) are considered. Both the 4- and 10-emitter cases contain three distinct radar measurements. In addition to bearing data for the multitarget case, the fuzzy clustering algorithm is used to cluster on RF and PRI, forming a RF-PRI sliding window. The bearing for each emitter is then associated with the radar data in the same way as in the single emitter cases. In Figs. 9-12, the ESM cluster that predominantly contains the emitter with constant bearing of 0° has been associated with the radar data. It is important to observe that both the fuzzy association algorithm and the TW algorithm use the RF-PRI fuzzy clustering results. Since the TW algorithm is not capable of deinterleaving data, it requires a preclustering operation, and the fuzzy clustering algorithm has been shown to be effective for deinterleaving [4].

In Figs. 7 and 8, the noiseless RF and PRI values are plotted for 4- and 10-emitter cases, respectively. The rectangles are centered at the noiseless values in each case and have sides of length 6 times the resolution limit of the simulated ESM system for both RF and PRI. The RF and PRI resolutions are measured relative to the appropriate axes in each case. In both Figs. 7 and 8, the emitter with a 0° bearing is confined to the box in the lower left corner centered at RF = 9400 and PRI = 1360. In some instances, two emitters with different bearings have the same RF and PRI. This accounts for the fact that there are three boxes in Fig. 7 for the 4-emitter case and 7 boxes in Fig. 8 for the 10-emitter case.

Figure 9 presents results for the radar example $\mu = 0^\circ, 1^\circ, -1^\circ$ and 4 emitters, as shown in Fig. 7. Since the radar results contain truth, i.e., a target moving with constant bearing of 0° , a good association algorithm will establish that there is a firm correlation between the ESM data and the 0° bearing track. Figure 9 plots the probability that the association algorithms establish a firm association between ESM data and the radar measurements. What distinguishes Figs. 1 and 9 is the number of emitters producing the data, 1 in Fig. 1 and 4 in Fig. 9. The data are preclustered in RF and PRI as points come in, only then does association take place. This class of simulation is more sophisticated in the sense that the initial clustering operation can fail, and points associated with the proper emitter can be neglected or points not associated with it included. The fuzzy association algorithm results are given by the curve marked with '+'s and the TW results are indicated by the unmarked continuous curve. The vertical axis indicates probability of firm correlation, and the horizontal axis indicates the number of data points necessary to establish that level of probability.

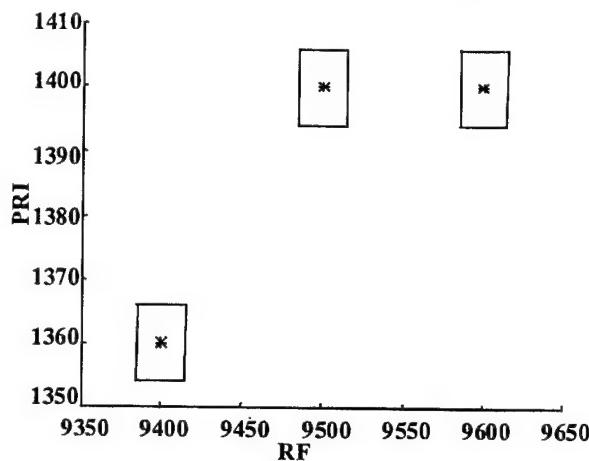


Fig. 7 — Four emitters RF and PRI with three standard deviation limits (boxes)

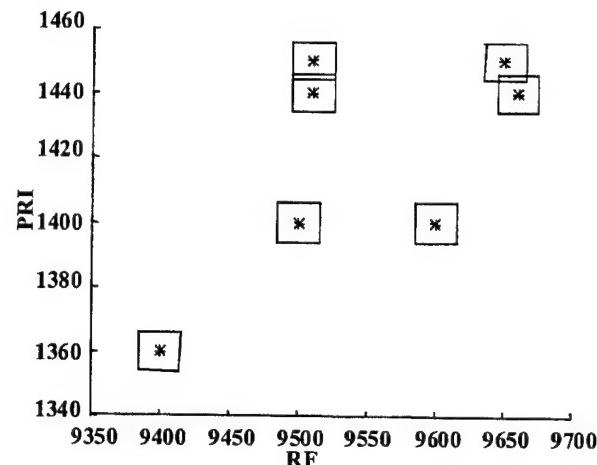


Fig. 8 — Ten emitters RF and PRI with three standard deviation limits (boxes)

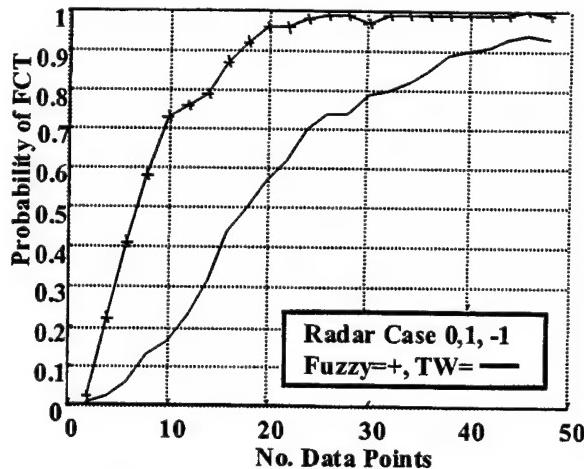


Fig. 9 — Noisy radar and four emitters for probability of FCT

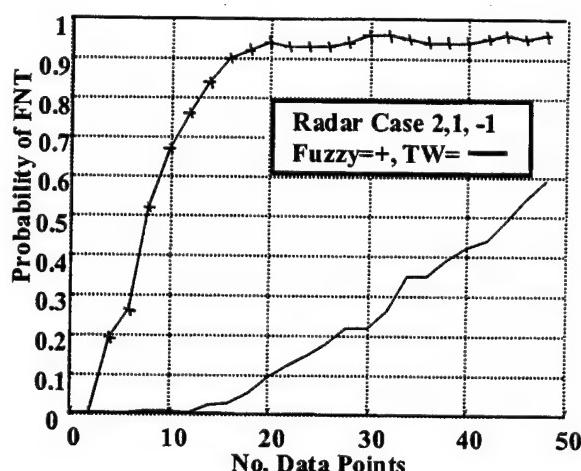


Fig. 10 — Noisy radar and four emitters for probability of FNT

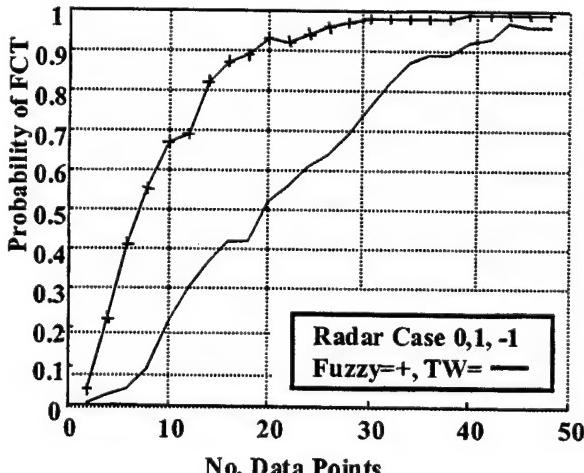


Fig. 11 — Noisy radar and 10 emitters for probability of FCT

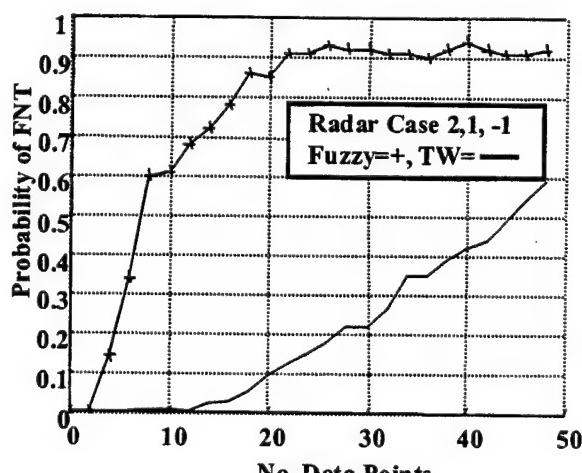


Fig. 12 — Noisy radar and 10 emitters for probability of FNT

The results for Fig. 9 are very close to those of Fig. 1; the fuzzy algorithm establishes high probabilities of FCT with 1/3 to 1/2 the data required by the TW algorithm. This is interesting since with 4 emitters there is an opportunity for bearing points to be misclustered in two ways. One kind of error that can arise is that bearing points from the wrong emitter are assigned to the emitter of interest, producing a significant outlier. The second kind of error relates to bearing points for the emitter of interest being assigned to another emitter, reducing the sampling rate. The first kind of error generally produces outliers and the second kind makes the data more intermittent. Probability augmented superclustering should be useful in dealing with both classes of errors. In this simulation, both the fuzzy association algorithm and the TW algorithm give performance similar to that of the single emitter case (Fig. 1). So it is not likely that errors of the first or second kind are dominating the calculation, as one would expect to see a rapid deterioration of the TW algorithm since it does not have outlier suppression or a significant ability to deal with intermittent data. The great similarity of results found in Figs. 1 and 9 suggest that the preclustering operation in RF and PRI was very successful in this case.

Figure 10 presents results for the radar example $\mu = 2^\circ, 1^\circ, -1^\circ$ and 4 emitters. Since the radar results do not contain truth, i.e., a target moving with constant bearing of 0° , a good association algorithm will establish that the ESM data are firmly not correlated with the radar tracks. There are 100 elements in this ensemble-average calculation.

Figure 10 plots the probability the association algorithms establish that the ESM and radar data are firmly not correlated. Once again the 4-emitter case is nearly indistinguishable from the 1-emitter case (Fig. 2).

Figure 11 provides results for the 10-emitter case for the radar example $\mu = 0^\circ, 1^\circ, -1^\circ$ with 0.1° standard deviation Gaussian random noise added. It is nearly indistinguishable from the 1- and 4-emitter cases, Figs. 3 and 9, respectively.

Figure 12 provides results for the 10-emitter case for the radar example $\mu = 2^\circ, 1^\circ, -1^\circ$ with 0.1° standard deviation Gaussian noise added. It is nearly indistinguishable from the 1- and 4-emitter cases, Figs. 4 and 10, respectively.

5.5 Multiple Closely Spaced Emitters

This section presents the results for 10 closely spaced emitters. In each case, there are three distinct radar measurements, $\mu = 0^\circ, 1^\circ, -1^\circ$. Close spacing of the emitters in the RF-PRI plane makes this a much more difficult problem than the one considered in section 5.4. In addition to bearing data for the multitarget case, the fuzzy clustering algorithm is used to cluster on RF and PRI, forming a RF-PRI window. The bearing for each emitter is then associated with the radar data in the same way as in the single emitter cases.

In Fig. 13, the noiseless RF and PRI values are plotted for the 10 emitters. The rectangles are centered at the noiseless values in each case and have sides of length equal to the resolution limit of the simulated ESM system for both RF and PRI. The RF and PRI resolutions are measured relative to the appropriate axes in each case. In Fig. 13, the emitter with a 0° bearing is confined to the box in the lower left corner centered at RF = 9400 and PRI = 1350.

In Figs. 14 and 15, the ESM cluster that predominantly contains the emitter with constant bearing of 0° has been associated with the radar data. It is important to observe that both the fuzzy association algorithm and the TW algorithm use the RF-PRI fuzzy clustering results. Since the TW algorithm is not capable of deinterleaving data, it requires a preclustering operation and the fuzzy clustering algorithm has been shown to be effective for deinterleaving [4].

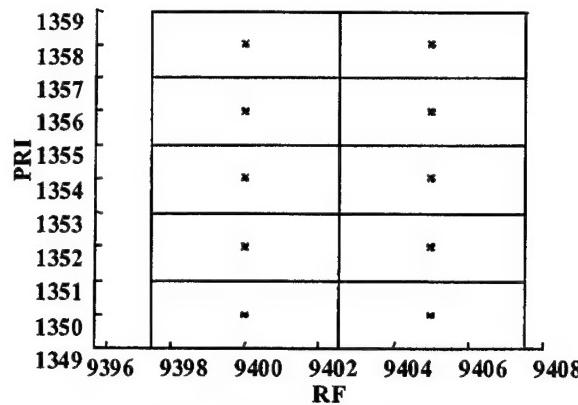


Fig. 13 — 10 closely spaced emitters

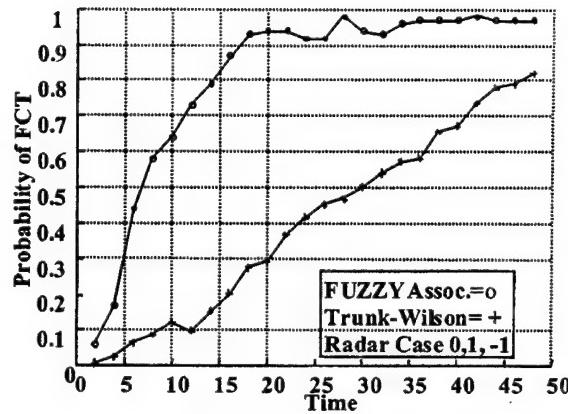
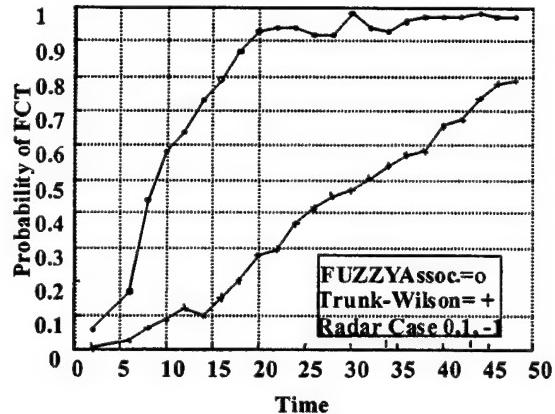
Fig. 14 — Noisy radar and 100% detection
for probability of FCTFig. 15 — Noisy radar and 90% detection
for probability of FCT

Figure 14 presents results for the radar example $\mu = 0^\circ, 1^\circ, -1^\circ$ and 10 emitters (Fig. 13). Since the radar results contain truth, i.e., a target moving with constant bearing of 0° , a good association algorithm will establish that there is a firm correlation between the ESM data and the 0° bearing track. Figure 14 plots the probability that the association algorithms establish a firm association between ESM data and the 0° radar measurement. The data are preclustered in RF and PRI as points come in; only then does association take place. As in the previous multiple-emitter examples, the initial clustering operation can fail, and points associated with the proper emitter can be neglected or points not associated with it included. The fuzzy association algorithm results are given by the curve marked with '+'s and the TW results are indicated by the unmarked continuous curve. The vertical axis indicates probability of firm correlation, and the horizontal axis indicates the number of data points necessary to establish that level of probability.

The results for Fig. 14 are fairly close to those of Fig. 3 for the fuzzy algorithm, which establishes high probabilities of FCT. This is interesting since with 10 emitters there is an opportunity for bearing points to be misclustered in two ways, as discussed in subsection 5.4. In this simulation, the fuzzy association algorithm gives performance similar to that of the single-emitter case (Fig. 3). The TW algorithm experienced significant deterioration, e.g., at 20 and 48 data points it dropped from 55 to 30% and 95 to 80%, respectively. It appears that the TW algorithm is sensitive to errors of the first and second kind. This is to be expected since the TW algorithm does not have the ability to deal with outliers and missed data that the fuzzy algorithm has.

In Fig. 15, the same 10-emitter calculation is considered as in Fig. 14 except the probability of detection of ESM points is 90%. In this case, the fuzzy algorithm reaches a 70% probability of FCT, between 3 and 4 times faster than the TW algorithm, and 90% probability of FCT at 18 data points, whereas the TW algorithm has a probability of FCT less than 80% for the first 48 points. The reduction in probability of detection of the ESM data by 10% has little effect on the fuzzy association algorithm and the TW algorithm.

In Fig. 16, the probability of detection of ESM points is 70%. The fuzzy association algorithm deteriorates little under the additional loss of data; its results are similar to those found in Figs. 3, 14, and 15. The TW algorithm has experienced significant deterioration. In Fig. 14 for 90% probability of detection, it has more than 80% probability of FCT at 48 data points; for a 70% detection rate its probability of FCT is less than 60%. Once again, the fuzzy association algorithm maintains its performance even though the TW algorithm experiences significant deterioration.

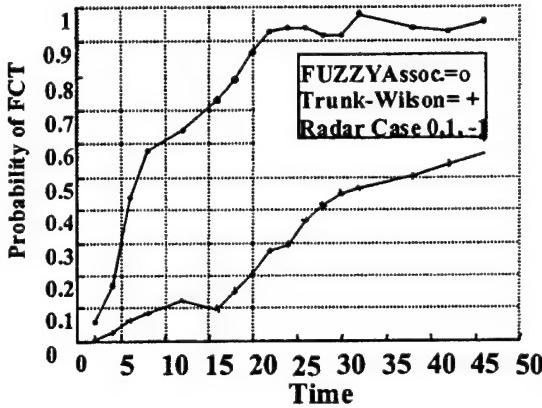


Fig. 16 — Noisy radar and 70% detection for probability of FCT

6. SLIDING WINDOWS, PERFORMANCE, AND COMPUTATIONAL COMPLEXITY

As discussed in section 2 and also in Refs. 7, 9-13, underlying the fuzzy association algorithm is fuzzy clustering. Roughly speaking the larger the number of points used in clustering, the better the algorithm's performance, because of improved cluster center estimation. There can be pathological cases where this is not true, but they are rare. Also, the algorithm's CPU time requirements tend to increase when larger data sets are clustered. A technique for reducing CPU requirements while maintaining performance is now explored.

The algorithm experiences a significant speed-up if a small amount of performance is sacrificed. One method of doing this is the introduction of a sliding window. If a window of size w is used, then for $n < w$ all n points are used in fuzzy clustering and superclustering operations. If $n \geq w$, then only w points are used. This results in a significant savings in CPU time and small loss in performance (Fig. 17).

Figure 17 is the probability of FCT vs the number of data points for the closely spaced 10-emitter example of Fig. 13. A sliding window in time has been applied to the ESM bearing data. This figure presents results for four different sliding window lengths. The sliding window lengths and the associated symbols they are marked with in the figure are 40 points (X), 30 points (*), 20 points (o), and 10 points (+). The fifth curve given in terms of dot-dashes represents the results for the TW algorithm for a 48-point window.

All four windowed fuzzy association calculations give performances superior to the TW algorithm. For a window of length w for $n < w$, the algorithm gives performance similar to the 48-point window of Fig. 14. For $n \geq w$, there is a slight deterioration in performance. The 30- and 40-point windows show the least performance loss and give results not far removed from those of Fig. 14.

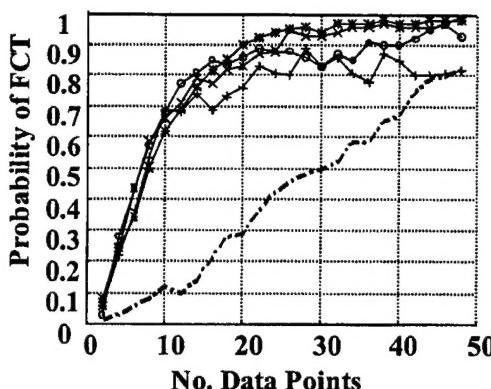


Fig. 17 — Sliding windows: 40 pts (X), 30 pts (*), 20 pts (o), 10 pts (+), and TW alg. (- - -)

The introduction of a sliding window results in a significant reduction in CPU time requirements for fuzzy clustering and superclustering. The maximum expected CPU times for the 10-, 20-, and 30-point windows are $1/5$, $1/2$, and $3/4$ of the maximum CPU time required by the 40-point window calculation. The effects of smaller windows and other approximations will be explored in the future.

Figure 18 compares the CPU time requirements for the four different windowed calculations. The vertical axis gives the ensemble-averaged CPU time usage (per data point for a Pentium 133) using MATLAB. The horizontal axis gives the number of data points. The curve markers 40 points (X), 30 points (*), 20 points (o), and 10 points (+) are the same as in Fig. 17.

Each time consumption curve takes the form

$$\tau(n, \omega) \equiv \begin{cases} 0.02n & n < \omega \\ 0.02\omega & n \geq \omega \end{cases}$$

where τ is the ensemble-averaged CPU time consumption per data point, n is the number of data points that have been detected, and ω is the window size. For $n \geq \omega$, the algorithm uses the same amount of time, on average, for clustering as would be expected since it has reached the window size. The linear rise for $n < \omega$ is harder to explain. It may indicate that the highly nonlinear recursive expressions of the Picard algorithm [5,6] could be linearized for this example.

This figure readily points up the effectiveness of windowing procedures. For $n \geq \omega$, CPU time consumption is proportional to window size. For the small loss in performance found in going from a 40-point window to a 10-point window, CPU time drops by a factor of 5.

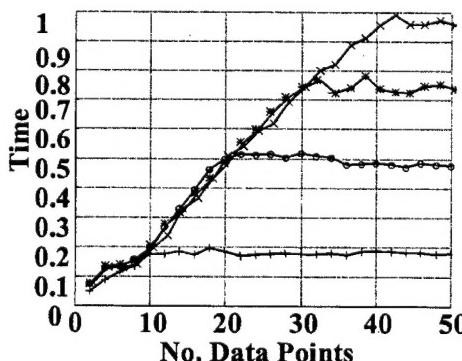


Fig. 18 — Average CPU consumption per data point. Sliding windows: 40 pts (X), 30 pts (*), 20 pts (o), and 10 pts (+)

All the calculations of Fig. 18 were carried out in MATLAB. A C-program would probably be 10 to 100 times faster. So for the 10-point window for $n \geq \omega$, the ensemble-averaged CPU usage per point would be between 0.002 and 0.02 s. The CPU-time estimates for a C-version of the software are probably conservative because an initial test of a C-version under development seems to indicate it is up to 250 times faster than the MATLAB version. Also, these tests were conducted on a Windows 95, 133 MHz PC. If the test had been conducted on a machine with a real-time operating system, i.e., one without interrupts, the CPU requirements would have been less.

With regard to the discussion above, the expression for CPU time consumption can also be written in the more compact form

$$\tau(n, \omega) \equiv kS_{CPU} \min(n, \omega),$$

where $S_{CPU} = 133$ MHz, $k = 1.5 \times 10^{-q}$ (1 sec/pt/MHz) with $q \equiv 4$ for MATLAB, and $5 \leq q \leq 7$ for the computer language C.

7. FUTURE APPLICATIONS

A resource allocation manager (RM) based on fuzzy logic is currently under development [14]. The RM will be used to control a multifunction antenna. The antenna provides ESM measurements, radar search, radar update, identification friend or foe (IFF), and communications. Some form of multisensor data association/correlation will be essential for the control of the multifunction antenna. The fuzzy association algorithm would be a very good choice for this application, especially when data are sparse, intermittent, and noisy. Also, since the RM will be based on fuzzy logic, linguistic data represented as fuzzy variables can be easily transferred between the RM and the fuzzy association algorithm. This will allow very efficient use of military information, i.e., expertise, intelligence, etc. The RM and related algorithms are described in more detail below.

The RM will allow codification of military expertise in a simple mathematical formalism known as the fuzzy decision tree [14]. The fuzzy decision trees will form what is known as a fuzzy linguistic description, i.e., a formal fuzzy if-then rule-based representation of the system. There will be two kinds of concepts on the fuzzy decision tree. The root concepts, i.e., the lowest and hence most primitive part of the tree, and the composite concepts. Composite concepts are built from root concepts. Since the decision tree is fuzzy the uncertainty inherent in the root concepts propagates throughout the tree, so the composite concepts are also intrinsically fuzzy.

The uncertainty in the root concepts will be given as fuzzy membership functions. The functional form of the fuzzy membership functions for the root concepts will be selected heuristically and will generally carry one or more free parameters. The free parameters in the root concepts will be determined by optimization, both initially and later at noncritical times.

A genetic algorithm will be used for optimization [15]. This is an algorithm based on an evolutionary paradigm. Simulated annealing [16], another stochastic optimization algorithm based on a thermodynamic model, may also be investigated.

There is also an intrinsic ID-fusion problem since many sources are capable of providing an ID declaration for a multifunction antenna of this type. How can these different identity declarations be combined to give an optimal identity declaration, automatically, i.e., without operator intervention? The ID-fusion algorithms that might be considered are of the following types: Bayesian, Dempster-Shafer, Fuzzy-Dempster-Shafer, and Generalized Evidence processing [17]. Time permitting, these algorithms will be compared and an optimal algorithm selected or invented.

The use of a fuzzy association algorithm in conjunction with a fuzzy RM and possibly a fuzzy ID algorithm will offer many advantages. The advantages would include high computational efficiency, easy codification of linguistic information including expertise, and a straightforward method of representing uncertainty and propagating it from one fuzzy algorithm to another.

8. CONCLUSIONS

A fuzzy logic algorithm for clustering and associating data measured on different sensors has been developed. It can estimate parameters such as bearing with only a small error. Two subcomponents of the algorithm, *superclustering* and *probability augmented superclustering*, can determine the number of targets present in the data.

The fuzzy algorithm's abilities as an association algorithm have been compared to the Trunk-Wilson (TW) association algorithm, a Bayesian philosophy algorithm. In simulations in which noiseless radar data contained truth, the fuzzy association algorithm establishes a firm correlation with 1/3 to 1/2 the data required by the TW algorithm. When the noiseless radar data did not contain truth, the fuzzy algorithm outperformed the TW algorithm with only 1/6 the data. When simulated radar data are used with noise of realistic magnitude, the fuzzy algorithm shows almost no deterioration in its performance. When the ESM and radar systems randomly fail to measure up to 30% of the data points, the fuzzy association algorithm shows little deterioration and is always better than the TW algorithm. The TW algorithm showed marked deterioration when data points are randomly lost or when there are many closely spaced emitters. The fuzzy algorithm exhibited almost no deterioration in performance when one to 10 emitters were present, even when there were many very closely spaced emitters. Introduction of a bearing sliding window can significantly reduce CPU requirements while offering little deterioration in performance. The fuzzy association algorithm's ability to make a correct decision with much less data than the TW algorithm is crucial since ESM data are generally sparse, intermittent, and noisy. Finally, the fuzzy association algorithm should be applicable to many different multisensor problems requiring high-quality decisions.

9. ACKNOWLEDGMENTS

This work was sponsored by the Office of Naval Research. The author gratefully acknowledges useful discussions with Mr. Ed Khoury and Dr. Joseph Lawrence III. The author also thanks Gay Polk, Sharon Davis, Ruth Smith, and Leroy Crispell for administrative help.

REFERENCES

1. G.V. Trunk and J.D. Wilson, "Association of DF Bearing Measurements with Radar Track," *IEEE Trans. AES*, **AES-23**(4) (1987).
2. H.J. Zimmerman, *Fuzzy Set Theory and Its Applications* (Kluwer Academic Publishers Group, Boston, 1991) p. 11.
3. H. Spath, *Cluster Analysis Algorithms for Data Reduction and Classification of Objects* (John Wiley and Sons, New York, 1980).
4. J.F. Smith III, "A Fuzzy Clustering and Superclustering Scheme for Extracting Structure from Data," NRL/FR/5740—96-9844, December 31, 1996.
5. J.C. Bezdek and J.C. Dunn, "Optimal Fuzzy Partitions: A Heuristic for Estimating the Parameters in a Mixture of Normal Distributions," *IEEE Trans. Comp.* **C-24**, pp 835-838 (1975).

6. J.C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms* (Plenum Press, New York, 1981).
7. J.F. Smith III, "A Fuzzy Logic Multisensor Association Algorithm," in *Signal Processing, Sensor Fusion, and Target Recognition VI*, I. Kadar, ed., Vol. 3068, pp. 76-87, *SPIE Proceedings*, Orlando, FL, April 21-24, 1997.
8. G.V. Trunk and J.O. Coleman, "Radar-ESM Correction Decision Procedure," NRL Report 8476, May 7, 1981.
9. J. Smith, A. Huynh, and M. Kim, "An Application of Clustering and Speed Discrimination to Tracking," NRL/MR/5740—95-9801, December 1995.
10. J.F. Smith III, "A Fuzzy Logic Multisensor Association Algorithm Dealing with Multiple Targets Intermittent Data and Noise," in *Signal Processing, Sensor Fusion, and Target Recognition VII*, I. Kadar, ed., Vol. 3374, *SPIE Proceedings*, Orlando, FL, April 13-17, 1998.
11. J.F. Smith III, "A Fuzzy Logic Multisensor Association Algorithm: Applied to Noisy, Intermittent and Sparse Data," *Fusion98, Proceedings of the International Conference on Multisource-Multisensor Information Fusion*, Las Vegas, Nevada, USA, July 6-9, 1998 (CSREA Press, Athens, Georgia).
12. J.F. Smith III, "A Fuzzy Logic Multisensor Association Algorithm: Theory and Simulation," NRL/FR/5740—97-9866, September 1997.
13. J.F. Smith III, "A General Fuzzy Logic Multisensor Association Algorithm, Theory and Application," *IEEE Trans. AES* (under IEEE review).
14. J.F. Smith III, "A Resource Allocation Manager for an Advanced Multifunction Aperture," prepared for NRL's Advanced Multifunction RF System (AMRFS) committee, 1999.
15. D.E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning* (Addison-Wesley, Reading, MA, 1989).
16. W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in Fortran: The Art of Scientific Computing* (Cambridge University Press, UK, 1992).
17. E. Waltz and J. Llinas, *Multisensor Data Fusion* (Artech House, Boston MA, 1990).